Efficient Content Distribution in an Information-centric Hybrid Mobile Network

Anders Lindgren
Swedish Institute of Computer Science
Box 1263
SE-164 29 Kista, Sweden
andersl@sics.se

***Abstract*—Information or content centric networking is believed by many to have great potential to be the appropriate networking paradigm for the future Internet. In information centric networking, focus is shifted from the end-points in the network to the information objects themselves, with less care being placed on from where the information is fetched. In addition to the benefits this networking paradigm has in fixed networks, it also simplifies operation in mobile networks and has the potential to improve performance.**

In this paper, we describe one way in which the NetInf network architecture can be used in a hybrid mobile network in an urban setting, and run simulations to evaluate the benefits that this approach can yield, both to the end users (in terms of improved performance such as reduced latency with over 50%), as well as to the operators in terms of a reduction of traffic load in the cellular access networks with up to 97%.

I. Introduction

Following the introduction of more and more smart-phones and other high-capacity devices using the cellular network, there has been a recent trend showing an explosive growth of data traffic in cellular networks [1]. This growth of data traffic usage in the cellular networks is problematic as operators get a high load on their networks and are forced to invest in more infrastructure. As flat-rate services for data traffic are becoming more common (and heavy data traffic users are the most likely ones to have such service), the increased network load does not provide increased revenue, but only increased costs. Another trend is however also that there are many locations where access points (APs) for WiFi or other local, high bandwidth, wireless network technologies are deployed. These APs usually provide a higher throughput than the cellular network, but do not have the same ubiquitous coverage as the 3G network. We envision using these access points, both as a way for mobile nodes to increase the speed of downloads, but also as locations where storage can be installed to allow for caching of popular content, further increasing the potential saving. Doing this it would be beneficial to the operators as the traffic load of the cellular network could be reduced, and end users would also gain as they would experience better performance.

This paper describes an information-centric networking approach to communication in such hybrid infrastructure/ opportunistic wireless networks. Simulations are performed in order to evaluate the use of such a paradigm in urban settings. We investigate the benefits it can provide both to the end users (in terms of improved performance such as reduced latency), as well as to the operators (e.g., in terms of reduced traffic load in the cellular access networks).

II. Information-centric Networking

Recently, many researchers have seen the potential in a switch of focus in computer networking from the previous host-centric view of networking in which the important data units are packets being sent between two hosts in the network to the new information-centric networking paradigm. In this paradigm, the focus is on the data and content that applications need instead of on the location of this data. One major benefit of such a networking paradigm (and the one mainly highlighted in this paper) are the inherent possibilites of caching of content in the network close to the end nodes in order to enhance end-user performance and reduce network utilization.

Multiple proposals exist for network architectures with information-centric approaches to networking, including PSIRP [2], Content-Centric Networking [3], [4] (CCN), and Network of Information[5] (NetInf). The NetInf architecture uses a name resolution service (which can be a combination of local and global lookup services) to map between content names and content locators (e.g., IP addresses). As content can be cached in the network, a name resolution can return multiple locators so that the node can select the "best" location of the data (according to some criteria, for example the closest copy of the content, or content located at the same local network). In this paper, we assume that there is a NetInf infrastructure in place in the network such that content can be cached at strategic locations (the WiFi access points in the simulations described in Section III), and that there is a name resolution system in place that will allow nodes to determine the best possible location from which to get the requested copy (in the simulations, it will be a matter of either getting a cached copy from the WiFi access point, downloading it from a location in the global network over the WiFi access point, or downloading it over the cellular network).

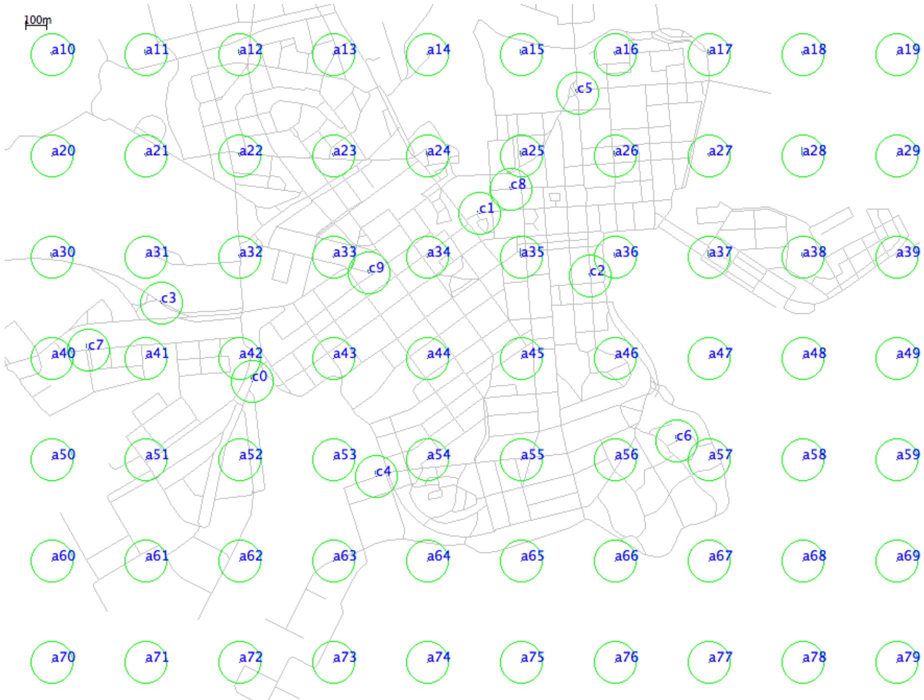

Fig. 1. Simulation scenario.

III. Simulation Setup

In this paper, we evaluate a scenario where mobile nodes in an urban area generate requests for some content (either content only relevant to them, or some popular content that a large portion of the nodes will request). If a node is not within reach of a WiFi AP at the time of the request, the download will start over the cellular 3G network. As soon as the node comes within range of an AP, it will continue the transfer on this network instead. As the content is downloaded to the AP, it is also cached there in order to serve it faster to future nodes that may request it at this AP. In order to further reduce the traffic in the cellular network, we also allow nodes to defer the download of content that is not time critical until some certain deadline. Nodes specify a defer deadline, which if set to zero means that the download should be initiated immediately, but if it is non-zero if means that if only the cellular network is available, download can be deferred until an AP is available or the deadline has passed. This means that content will be downloaded at most "defer deadline" seconds later than if that mechanism had not been used, but the traffic in the cellular network can be further reduced.

The urban scenario described above was simulated using a model of vehicular traffic in the downtown Helsinki area. This model is available in the ONE simulator for DTN networks[6] and we used this to generate the mobility patterns in the simulation. The network simulations were done in a custom simulator written for this purpose. WiFi access points were distributed in a uniform grid pattern at different densities for different simulation runs. The number of access points vary from none (nodes only use the cellular network) up to 374 APs (which corresponds to approximately 77% of the simulation area being covered by WiFi access points). 500 cars move around on the streets of Helsinki, choosing routes based on certain predefined points of interest in the city. Figure 1 shows an example of the simulated area. During the simulation, the cars (the mobile client nodes) issue requests for different information objects of varying size and popularity (ranging from each piece of content only being requested by a single node, to the content being requested by all mobile nodes) from the cellular network. We assume that requests are small in comparison to the size of the requested content. If the data is delay-tolerant, the nodes can also set different acceptable deadlines for the requested information objects in order to further reduce the traffic over the cellular network. This is done through the node specifying a future time when it at the latest want to have the data delivered. The node estimates the time it would take to download the entire information object over the cellular network, and as long as there is more time than this until the delivery deadline, it defers the download and only downloads the information object when in range of a WiFi access point, thus reducing load on the cellular network. If the time to the delivery deadline is less than or equal to the estimated download time over the cellular network, the download is initiated over the cellular network in order to ensure (with high probability) that the data is delivered before the deadline. Table I shows the simulation parameters and the different values used.

The name resolution system and locator selection works in the following preference order. If a mobile node is within range of a WiFi access point (AP) and the requested content is cached here, the cached copy of the content will be downloaded by the mobile node. If it is in range of an AP that does not have the content cached, it will begin the download of the content over the AP from a location in the global network. As the content is downloaded through this path, it is also cached at the AP to serve future requests. If the mobile node is out of range from an AP, it will initiate the download of the content over the cellular network from a location in the global network. If a node that has started a download over the cellular network later comes into contact with an AP, it can switch access method and continue the download through the AP (either from a cached copy of from the global network).

IV. Results

Initially, we consider simulations where no deadline was specified for the content (i.e., nodes will try to download it as quickly as possible). In Figure 2, we can see how the download time to the end user for different file size varies as an increasing number of WiFi access points (APs) are deployed into the system. The figure clearly shows a potential for large performance gains through the use of a hybrid system like this. It is also clear that the benefit is greater for larger file sizes, as small file transfers that are initiated outside the range of an AP are more likely to finish before reaching an AP.

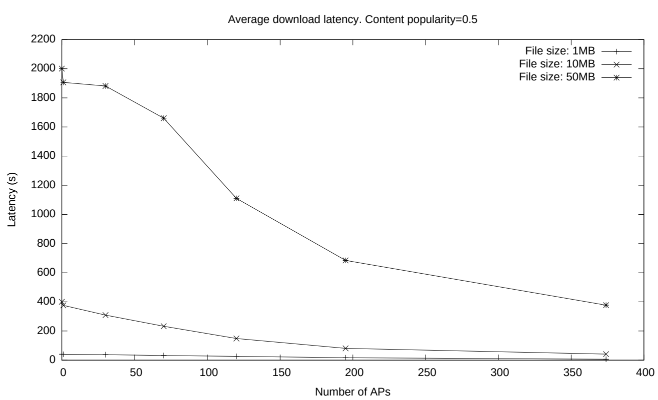

Fig. 2. Content download latency for different number of APs and file sizes. Delayed deadline=0 (immediate transmission). Content popularity=0.5.

In Figure 3, we study the impact the deployment of the access points has on the fraction of data that is delivered through the cellular network. As with the delay, the impact is smaller when the content is very small, but we can see similar reductions for the larger files sizes. For the larger file sizes, deploying 70 APs (corresponding to covering as little as 14% of the simulated area) reduces the traffic in the cellular network with more than 50%, and by deploying more APs, the traffic in the cellular network can be reduced with up to 97%.

Figure 4 shows effect of the defer deadline mechanism on the fraction of content delivered through the cellular network for some different number of APs. While waiting for a long time (up to one hour) reduces the traffic through the cellular

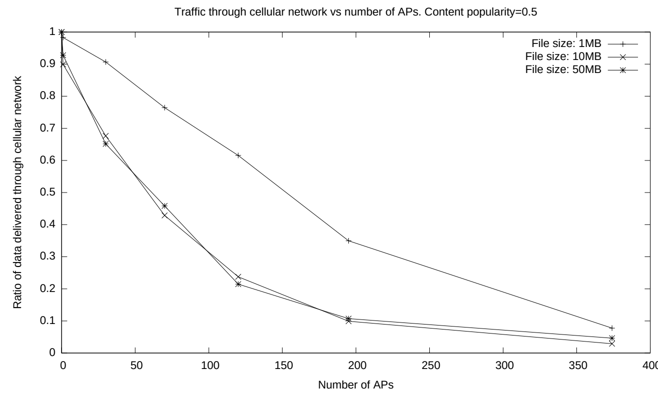

Fig. 3. Traffic through the cellular network for different number of APs and file sizes. Shown as a fraction of the traffic sent through the cellular network using NetInf, in comparison to the amount of traffic sent when only using the cellular network. Delayed deadline=0 (immediate transmission). Content popularity=0.5.

network to a very low number, deferring transmission for this long might not be desirable in most situations. However, being able to defer the transmission for 5-10 minutes already reduce the traffic with up to 4 times. This is a very reasonable amount of time to expect people to be willing to defer downloads, for example still ensuring that the content is available as they reach their destination.

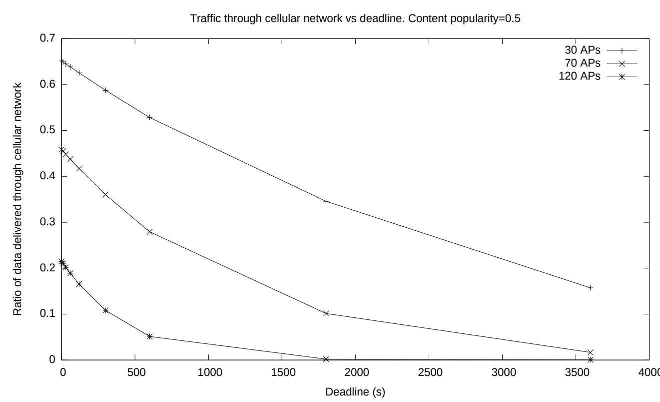

Fig. 4. Traffic through the cellular network for different delayed deadlines and for different number of APs. Content popularity=0.5.

The previous two figures have shown how much of the traffic flows through the cellular network as more APs are added. As the traffic diverted from the cellular network is sent over WiFi APs instead, more traffic is generated in the access network to these APs. While this resource is likely to be less scarce than that of the cellular network, it is still of interest to minimize the resource usage. One major benefit that NetInf has over other multi-access schemes that try to combine cellular and WiFi access is the inherent caching of popular content made possible by adding storage in the network. In Figure 5, we show the amount of traffic that is sent through the WiFi access network for different levels of content popularity (here

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Number of WiFi APs	0, 1, 30, 70, 120, 195, 374
Number of cars	500
Car speed	10-50 km/h
Simulation time	12 hours
WiFi bandwidth/goodput	4 Mbit/s
3G bandwidth/goodput[7]	200 kbit/s
AP access network bandwidth	10 Mbit/s
File sizes	1, 10, 50 MB
Content popularity	0.1 - 1.0
Deadline (deferred transmission)	0, 10, 30, 60, 120, 300, 600, 1800, 3600 s

shown as the number of nodes that eventually request a given information object). When only a single node requests each object, all data transferred must go through either the cellular or the WiFi access network. As content become more popular and more nodes start requesting it, we can see the effects of being able to cache the content at the WiFi access points as the traffic sent through the access network drops. There is a rather drastic decrease in traffic through the access network as we see an initial increase in the content popularity, with traffic reductions of over 50% when as few as 10 nodes (2% of the total number of nodes) request a certain piece of content.

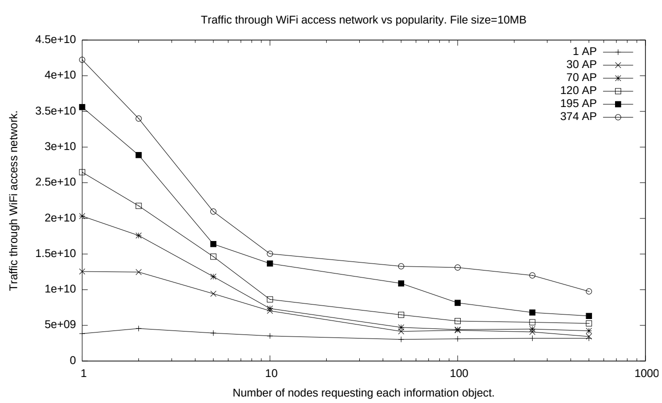

Fig. 5. Traffic through the WiFi access network for different content popularity (number of nodes requesting each information object) and for different number of APs. File size=10MB. The x axis is shown in log scale to make the graph easier to read

The results in the figures above show that there are large gains from using NetInf in this type of scenario. The users will benefit from this as the total download latency will be lower for time critical data. The operators on the other hand will be able to greatly reduce the traffic through their cellular networks (with up to 97%), especially if data is somewhat delay tolerant so that transmission can be deferred until an AP is encountered. There exist other multi-access schemes that can achieve some of the same benefits by switching between the cellular network and other networks, such as the WiFi access points. The benefits of using NetInf over these other mechanisms are however twofold. First of all, the total amount of traffic can be reduced through the caching in NetInf that helps reduce the traffic through the WiFi access network. Secondly, when using NetInf in the network, there is no need for additional mechanisms (such as mobile IP, seamless handovers at layer 2/3, etc) to do the multi-access, but the inherent properties of NetInf that locates the best copy of an information object allows this to happen automatically.

V. RELATED WORK

There exist other papers that have investigated hybrid mobile networks in urban scenarios, as well as different delay-tolerant networking approaches to these. In [8], a system for web access over WiFi hotspots in combination with a DTN-type of network where mobile nodes forward data is designed. This paper use the same city mobility model as in our evaluations, but use pedestrians instead of vehicles for the mobile nodes. Here, the focus is not on the information-centric networking concept, but rather on the web specific architecture. Further, it differs from this paper as it only uses the WiFi access points and opportunistic contacts between mobile nodes for data delivery and does not make use of the cellular network.

In two papers by Hui and Lindgren [9], [10], the problem of using an infrastructure of WiFi access points in conjunction with opportunistic communication between mobile nodes is studied for a couple of different urban scenarios. These papers differs from this paper, as they have a more theoretical focus, in which they try to find the upper bound for what can be achived in these scenarios.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we have shown some of the benefits of applying a information-centric communication paradigm in a hybrid mobile network with a combination of cellular connectivity and opportunistic contacts with high-bandwidth access points. We show that this approach is appealing both to the end-users as well as the network operators. The end-users benefit from this through improved performance in terms of reduced download latencies (especially for large file sizes) when content can be access from cached copies in nearby locations or downloaded over an alternative network connection (the WiFi network). In the simulations in this paper, download latency was reduced with more than 50%, which should be a strong enough incentive for a users to want to use this system. The operators on the other hand are able to greatly reduce the amount of traffic in the cellular networks, and for popular content also the traffic in the access networks of the WiFi access points. These capacity savings are shown to be even greater (up to a 97% reduction in cellular network load) when data is delay tolerant so that downloads can be deferred for a while in order to hopefully be able to be downloaded without using the cellular network. All this functionality could be implemented in the current Internet architecture without the use of NetInf or any other information-centric networking architecture. The presence of such an architecture however takes this mode of operation from a special case for which lots of extra protocols are needed (for seamless handovers, cache management, etc) to the normal mode of operation that is inherently supported by the network infrastructure. Thus, we see a great potential for the use of this type of network architecture in mobile networks.

Performance could potentially be further enhanced if caching at the mobile nodes is also permitted, such that mobile peers can exchange data as they meet. This does however introduce additional complexity and gaining user acceptance might be difficult due to the risk that such a mode of operation would drain their batteries quickly. Thus, we do not include this option in the evaluations in this paper, but will consider it (along with appropriate measures to prevent unfair resource usage) in future work. In addition, it is also of interest to study other AP deployment patterns than the uniform grid pattern used in this paper. Access points could be deployed at intersections with heavy traffic or at pre-defined points of interest (which is available in the city model) first in order to further enhance the performance improvements. As these improvements to the simulation model only should make the benefits of an information-centric architecture even more apparent, we believe that all the conclusions drawn in this paper would still be valid.

VII. Acknowledgements

This work has been supported by the IST 7th Framework Programme Integrated Project 4WARD, which is partially funded by the Commission of the European Union.

References

[1] "Mobile data traffic explosion tops mobile and wireless industry trends for 2010," The Independent, UK, January 3 2010.

[2] M. Ain *et al.*, "Architecture definition, component descriptions, and requirements," PSIRP EU FP7 Project, Deliverable D2.3, Feb. 2008, fP7-INFSO-ICT-216173-PSIRP-D2.3.

[3] V. Jacobson, M. Mosko, D. Smetters, and J. Garcia-Luna-Aceves, "Content-centric networking," Palo Alto Research Center, Whitepaper, Jan. 2007.

[4] C. Partridge and V. Jacobson, "Making the case for content-centric networking," *ACM Queue*, vol. January, pp. 8–16, 2009.

[5] B. Ahlgren, M. D'Ambrosio, C. Dannewitz, M. Marchisio, I. Marsh, B. Ohlman, K. Pentikousis, R. Rembarz, O. Strandberg, and V. Vercellone, "Design considerations for a network of information," in *Proceedings of the First International Workshop on Re-Architecting the Internet (ReArch2008)*, Madrid, Spain, December 2008.

[6] A. Keränen, J. Ott, and T. Kärkkäinen, "The ONE Simulator for DTN Protocol Evaluation," in *SIMUTools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*. New York, NY, USA: ICST, 2009.

[7] K.Pentikousis, M. Palola, M. Jurvansuu, and P. Perala, "Active goodput measurements from a public 3g/umts network," *IEEE Communications Letters*, pp. 802–804, Sep. 2005.

[8] M. Pitkänen, T. Kärkkäinen, and J. Ott, "Opportunistic web access via wlan hotspots," in *Proceedings of the 8th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom)*, March 2010.

[9] P. Hui, A. Lindgren, and J. Crowcroft, "Empirical evaluation of hybrid opportunistic networks," in *Proceedings of the First International Conference on COMmunication Systems and NETworkS (COMSNETS 2009)*, January 2009.

[10] P. Hui and A. Lindgren, "Phase transitions of opportunistic networking," in *Proceedings of the ACM SIGMOBILE workshop on Challenged Networks (CHANTS 2008)*, September 2008.